Fuzzy Systems and Data Mining IX
A.J. Tallón-Ballesteros and R. Beltrán-Barba (Eds.)

doi:10.3233/FAIA231042

Concentration Diagnosis in Soft Sensing Based on Bayesian T-Distribution Mixture Regression

Changhai Xia [a,1], Zhiping Peng [b] , Delong Cui [c] , Qirui Li[c], and Lihui Sun[a]
[a] *School of Information and Control Engineering, Jilin Institute of Chemical Technology, Jilin 132022, China*
[b] *Jiangmen Polytechnic College, Jiangmen City 529090, China*
[c] *School of Computer Science, Guangdong University of Petrochemical Technology, Maoming City 525000, China*

Abstract. This study presents a cutting-edge soft sensing approach for coke-making diagnostics, aimed at tackling the challenges posed by multifaceted, nonlinear, non-Gaussian, and noisy operational data prevalent in coke-making ovens. Our proposed method leverages a Bayesian t-distributed mixed regression model, effectively capturing the intricate nature of multivariate, nonlinear, and non-Gaussian data. The utilization of the t-distribution ensures the model's resilience to interference, with model parameter estimation achieved within a Bayesian framework. Conducting simulation experiments and real industrial experiments, as well as comparative analysis with PLSR, GMR, and GPR models, we demonstrate the model's good robustness, excellent prediction accuracy, and robustness, further confirming its potential application in coking diagnosis.

Keywords. Steam Cracking; t-distribution; t-distribution mixture regression model; Variational Bayesian; Robustness

1. Introduction

An ethylene cracking furnace plays a pivotal role in steam cracking production[1]. The production capacity and technology of the cracking furnace directly influence the production scale, yield, and product quality of the entire ethylene plant[2]. Due to the unique characteristics of hydrocarbon cracking feedstock, the cracking reaction at high temperatures inevitably leads to the formation of coke particles, which adhere to the inner wall of the furnace tube, a phenomenon known as furnace tube coking[3]. Furnace tube

[1] Corresponding Author, Changhai Xia, Jilin Institute of Chemical Technology, Jilin 132022, China; E-mail: xch1195940@163.com.

The work presented in this paper was supported by: National Natural Science Foundation of China (62273109, 61772145, 61672174), the Guangdong Basic and Applied Basic Research Foundation (2021A1515012252, 2020A1515010727, 2022A1515012022, 2019A1515010830, 2023A1515011913), the Key Field Special Project of the Department of Education of Guangdong Province (2020ZDZX3053), the Key Realm R&D Program of Guangdong Province (2021B0707010003), and the Maoming Science and Technology Project (210429094551175, 2022DZXHT028, and mmkj2020033), 2022 Guangdong Province Science and Technology Innovation Strategy Special Project (pdjh2022b0349), Projects of PhDs' Start-up Research of GDUPT (2023bsqd1002).

coking poses a significant threat to ethylene production and can even result in catastrophic accidents in severe cases[4-5]. Consequently, accurately predicting the extent of cracker tube coking holds the utmost importance.

Currently, two main methods are used for tube coking diagnosis in ethylene cracking furnaces: coking diagnosis based on mechanistic models[6-7] and data-driven coking diagnosis through soft sensing[8]. Although mechanistic model-based approaches have been employed early on, their accuracy in coking inference is limited due to challenges in accurately obtaining parameters for crucial mechanistic models[9-11]. On the other hand, data-driven coking diagnosis offers a promising solution by utilizing available cracking process parameters as input variables for artificial intelligence algorithms, establishing relationships with coking thickness through feature extraction and modeling[12]. Among these techniques, artificial neural networks have been widely adopted for identifying operating conditions and developing a system model for the stochastic distribution of outlet temperatures, laying the foundation for advanced coke on temperature (COT) control during the ethylene cracking furnace coking process[13-14]. Given the multi-modality, nonlinearity, and non-Gaussian characteristics of cracking furnace data, modeling methods based on mixture models have proven advantageous, with the Gaussian mixture model (GMM) being commonly employed[15]. Building upon GMM, the Gaussian mixture regression model (GMR) enables predictions of output variables[16-17]. However, in practical scenarios, operating data are susceptible to noise, and the use of Gaussian distribution, with its short tails, may lead to reduced model accuracy[18]. In contrast, the t-distribution mixture model, characterized by wider tails, exhibits better robustness to outliers[19].

To overcome the challenges posed by multi-modality, nonlinearity, non-Gaussian, and strong noise in cracking furnace operating data, this paper introduces a soft sensing method for tube coking diagnosis called the Bayesian t-distribution mixture regression model (STMR). The proposed approach effectively captures the complex features of the multi-modal, nonlinear, and non-Gaussian data by employing the mixture model. By utilizing the t-distribution, the model ensures robustness and accuracy, even in the presence of noise. Parameter estimation for the model is achieved through the variational Bayesian method, enhancing the reliability of the results. Finally, the predictive accuracy of the model is validated through experiments using real-world industrial data, demonstrating its practical utility and effectiveness.

2. Bayesian t-distribution mixture regression model

2.1. Bayesian t-distribution mixture regression

Given the auxiliary variable $\mathrm{X} = \{x_n\}_{n=1}^{N}$, and the dominant variable $\mathrm{Y} = \{y_n\}_{n=1}^{N}$, where N is the number of samples in the dataset, and the data dimensions for the auxiliary variables denoted by *d*, assuming that auxiliary variable X within the input space is a mixture of M t-distributions

$$p(x_n \mid \mu, \Lambda, v, \pi) = \sum_{m=1}^{M} \pi_m \mathcal{St}(x_n \mid \mu_m, \Lambda_m, v_m) \tag{1}$$

where μ is the mean, Λ is the precision, v is the degree of freedom, M is the number of mixed membership components, $\pi = \{\pi_m\}_{m=1}^{M}$ is the mixing coefficients also known as weighting coefficients, where the element π_m represents the a priori probability value of the mth membership component. The constraints are $0 \le \pi_m \le 1$, $\sum_{m=1}^{M} \pi_m = 1$.

It is assumed that the auxiliary variable X and the dominant variable Y under each mixed t-distribution component in the output space obey a linear relationship, i.e.

$$y_n = x_n^T \omega_m + \varepsilon_m \tag{2}$$

where ω_m is the regression coefficient between y_n and x_n, ε_m is the measurement noise, and $\varepsilon_m \sim \mathcal{N}(0, \sigma_m)$.

We can be obtained

$$p(y_n \mid x_n, \omega, \sigma) = \prod_{m=1}^{M} \mathcal{N}(y_n \mid x_n^T \omega_m, \sigma_m) \tag{3}$$

where $\omega = \{\omega_1, \omega_2, \ldots, \omega_m\}$ and $\sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_m\}$.

To facilitate variational processing, discrete binary indicator hidden variables $Z = \{z_n\}_{n=1}^{N}$ and robust hidden variables $U = \{u_n\}_{n=1}^{N}$ are introduced for x_n. where $z_{nm} \in \{0,1\}$, $\sum_{m=1}^{M} z_{nm} = 1$, and z_{nm} denotes the indicator that x_n comes from the mth mixing component, which is indicated when $z_{nm} = 1$, and which is assigned to this x_n, and vice versa when $z_{nm} = 0$. In particular, note that each sample can be for only one mixing component, and the mixing coefficient π is given to characterize the weight that each t-distribution component carries in the mixing model, and thus the conditional probability density function of the indicator hidden variable Z and the robust hidden variable U can be expressed as

$$p(z|\pi) = \prod_{n=1}^{N} \prod_{m=1}^{M} \pi_m^{z_{nm}} \tag{4}$$

$$p(z|\pi) = \prod_{n=1}^{N} \prod_{m=1}^{M} \mathcal{G}\left(u_{nm} \middle| \frac{v_m}{2}, \frac{v_m}{2}\right)^{z_{nm}} \tag{5}$$

Using the conjugate prior property of the exponential family of distributions, the prior distributions of the model variables are chosen in the following manner:

The prior distributions for the mixture coefficients π follow a Dirichlet distribution $\mathcal{D}ir(\pi|k_0)$, with k_0 being the hyperparameter of the Dirichlet prior. The precisions Λ_m of each mixture component follow the Wishart prior distribution $\mathcal{W}(\Lambda_m|r_0, s_0)$, where r_0 and s_0 are the degrees of freedom and scale matrix of the Wishart prior, respectively. The means μ_m of each mixture component follow the Gaussian distribution $\mathcal{N}(\mu_m|m_0, \rho_0)$, with m_0 and ρ_0 being the mean and precision parameters of the Gaussian prior, respectively. The coefficients ω_m in the linear relationship between the auxiliary variables X and dominant variables Y are drawn from the Gaussian distribution $\mathcal{N}(\omega_m|0, \lambda_m^{-1}.I)$, where $\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$, and λ_m represents the precision

parameter of the Gaussian prior for ω_m. As both λ and σ are represented as the precisions of Gaussian distributions in the model, the conjugate prior distributions for λ_m and σ_m are specified as Gamma distributions, with a_0, b_0, c_0 and d_0 as the hyperparameters.

The degrees of freedom parameters $v = \{v_m\}_{m=1}^{M}$ have no prior distributions and can be estimated by maximizing the lower bound on v_m.

The joint distribution hierarchy between the data and the variables is expressed as

$$\begin{aligned} p(Y, X, Z, U, \pi, \mu, \Lambda, \omega, \lambda, \sigma) = p(Y \mid X, Z, \omega, \sigma) p(X \mid Z, U, \mu, \Lambda) \\ \times p(U \mid Z) p(Z \mid \pi) p(\pi) p(\Lambda) p(\mu) p(\omega \mid \lambda) p(\lambda) p(\sigma) \end{aligned} \tag{6}$$

2.2. Variational inference for Bayesian t-distributed mixed regression

In the Bayesian framework, variational inference is adopted to estimate the model parameters. The variable $\Theta = \{\Theta^l, \Theta^{vb}\}$ is set to be the ensemble of the latent variable $\Theta^l = \{Z, U\}$ and the variational Bayesian variable $\Theta^{vb} = \{\pi, \mu, \Lambda, \omega, \lambda, \sigma\}$. Considering the approximate posterior $q(\Theta)$, the logarithm of the likelihood can be expressed as

$$ln\, p(X, Y) = \mathcal{L}(q) + KL[q(\Theta)||p(\Theta|X, Y)]\, ln\, p(X, Y)] \tag{7}$$

where $KL[q(\Theta)||p(\Theta|X, Y)]$ is the relative entropy in information theory and satisfies $KL[q(\Theta)||p(\Theta|X, Y)] \geq 0$. It is called KL scatter, which denotes the distance between the true posterior distribution $p(\Theta|X, Y)$ and the approximate posterior $q(\Theta)$. $KL[q(\Theta)||p(\Theta|X, Y)] = 0$ if and only if $p(\Theta|X, Y) = q(\Theta)$. so there is $\mathcal{L}(q) \leq ln\, p(X, Y)$, and $\mathcal{L}(q)$ is an evidential lower bound for $ln\, p(X, Y)$. The variational a posteriori for the variable Θ can be obtained by taking the variational components of $\mathcal{L}(q)$ for each variable in turn.

The hidden variables of the variable Θ are independent of each other and of the individual hyperparameters, factorizing the posterior distribution of the variates, i.e:

$$q(\Theta) = q(Z, U, \pi, \mu, \Lambda, \omega, \lambda, \sigma) = q(Z)q(U)q(\pi)q(\mu)q(\Lambda)q(\omega)q(\lambda)q(\sigma) \tag{8}$$

2.3. Parameter estimates

The approximate posterior distribution of the latent variable Θ^l is computed below

$$q(Z) \propto exp \sum_{n=1}^{N} \sum_{m=1}^{M} {z_{nm}}^{r_{nm}} \tag{9}$$

Consider the uncertainty on the robust variables. After normalization, the expectation of the indicator variable Z is

$$\langle z_{nm} \rangle = \frac{r_{nm}}{\sum_{m=1}^{M} r_{nm}} \tag{10}$$

The posterior distribution $q(U) = \mathcal{G}(u_{nm}|\alpha_{nm}, \beta_{nm})$ of the robust variable U, where

$$\alpha_{nm} = \frac{v_m + \langle z_{nm} \rangle d}{2} \tag{11}$$

$$\beta_{nm} = \frac{v_m + \langle z_{nm} \rangle \langle (x_n - \mu_m)^T \Lambda_m (x_n - \mu_n) \rangle}{2} \tag{12}$$

In addition, we have $\langle u_{nm} \rangle = \alpha_{nm}/\beta_{nm}$, $\langle Ln\, u_{nm} \rangle = \psi(\alpha_{nm}) - ln\, \beta_{nm}$, where $\psi(\cdot)$ is the digamma function.

Similarly, derive the posterior distributions of the other model parameters. In addition, the degree of freedom parameter v_m can be optimized by maximizing the log-likelihood function for v_m, since there is no information about the prior distribution.

In order to simplify the numerical computation, in this paper, the nonlinear equations are not solved, and using Stirling's formula $ln\ \Gamma(z) \approx (z - 1/2) ln\ z - z + 1/2 ln\ 2\pi + O(1/z)$, the updating formula for v_m can be obtained

$$v_m = -\left(1 + \frac{\sum_{n=1}^{N} \langle z_{nm} \rangle \left(\langle ln\, u_{nm} \rangle - \langle u_{nm} \rangle \right)}{\sum_{n=1}^{N} \langle z_{nm} \rangle}\right)^{-1} \tag{13}$$

The purpose of the parameter update is to maximize the evidence lower bound $\mathcal{L}(\mathrm{q})$, which is computed below based on the variational posterior of the parameter

$$\mathcal{L}(q) = \langle ln\, p(X, Y, Z, U, \pi, \mu, \Lambda, \omega, \lambda, \sigma) \rangle - \langle ln\, q(Z, U, \pi, \mu, \Lambda, \omega, \lambda, \sigma) \rangle \tag{14}$$

The difference of the evidence lower bound $\mathcal{L}(q)$ is computed after each iteration, and the algorithm is recognized to have converged when this difference is below a predetermined queer value ξ. i.e:

$$|\mathcal{L}(q)_t - \mathcal{L}(q)_{t-1}| < \xi \tag{15}$$

3. Soft measurements based on STMR

The data of the ethylene cracker consists of two parts: input data x includes the entrance and exit temperatures of the furnace tube, the outer surface temperature of the furnace tube, the adiabatic pressure ratio, the pressure across the section, and the venturi pressure, and the output data y is the coking degree of the furnace tube. Assuming that the new input variable is x_{new} and the corresponding output variable is y_{new}, in order to predict the output, it is necessary to construct a soft-measurement model based on Bayesian t-distribution mixture regression. For the mth mixture component, the posterior probability about x_{new} and the conditional probability about y_{new} are estimated as

$$p(z_{new} = 1 | x_{new}) = \frac{\langle \pi_m \rangle \mathcal{St}(x_{new} | \langle \mu_m \rangle, \langle \Lambda_m \rangle, v_m)}{\sum_{m=1}^{M} \langle \pi_m \rangle \mathcal{St}(x_{new} | \langle \mu_m \rangle, \langle \Lambda_m \rangle, v_m)} = S_{nm} \tag{16}$$

$$p(y_{new} | x_{new}) = \sum_{m=1}^{M} S_{nm} \mathcal{N}\left(y_{new} | x_{new}^T \langle \omega_m \rangle, \langle \sigma_m^{-1} \rangle + x_{new}^T \varphi_m^{-1} x_{new} \right) \tag{17}$$

Finally, the predicted output $\hat{y}_{new}$ is

$$\hat{y}_{new} = \sum_{m=1}^{M} S_{nm} x_{new}^{T} \tau_m \tag{18}$$

4. Experiments and analysis of results

Within this section, a numerical simulation and an application to an actual industrial process are employed to authenticate the viability and efficacy of the soft measurement algorithm based on STMR. In parallel, the Gaussian mixture regression model (GMR), Gaussian process regression model (GPR), and partial least squares regression (PLSR) are chosen for comparative analysis. For GMR, the optimal number of components is determined using the Bayesian Information Criterion (BIC). In tandem, the root mean square error (RMSE) is employed in this study to assess the performance of various models. The configurations of the used computer are given as follows: CPU: Core i7-1165G7@2.80GHz, RAM: 16 GB, OS: Windows 10, and Software: MATLAB (R2016a).

4.1. Numerical simulation

A system is set up with a 2-dimensional input variable $X = (x_1, x_2)^{\mathrm{T}}$ and an output variable Y. The input variable x obeys a mixed Gaussian distribution, and the relationship between X and Y is shown in Eq. (2). The specific parameter settings of the model are shown in Table 1.

Table 1 Parameters of three Gaussian components for numerical simulation

	m=1	m=2	m=3
π_m	0.2	0.2	0.6
μ_m	$(-3,1)^{\mathrm{T}}$	$(2,4)^{\mathrm{T}}$	$(3,-3)^{\mathrm{T}}$
Λ_m	$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$	$\begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}$
ω_m	$(1,1,-2)^{\mathrm{T}}$	$(1,-1,0)^{\mathrm{T}}$	$(-1,1,1)^{\mathrm{T}}$
σ_m	0.25	0.25	0.25

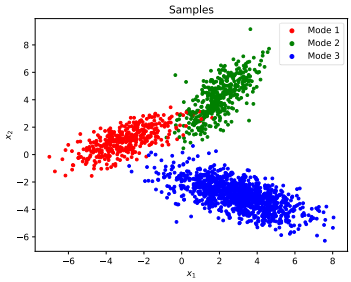

Figure 1 Spatial distribution of pure data

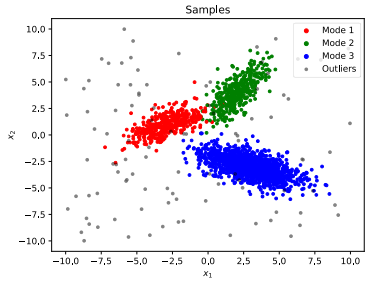

Figure 2 Spatial distribution of data under 7% noise pollution

The overall dataset comprises 5000 samples, with 2000 samples allocated to the training set and 3000 samples assigned to the prediction set. To assess the model's robustness, the training set is intentionally subjected to noise pollution by randomly introducing 1%, 5%, and 7% outliers. These outliers simulate adverse conditions and contribute to the dataset's noise-contaminated variant. Figure 1 visually portrays the spatial distribution of pristine data untarnished by noise, while Figure 2 illustrates the spatial distribution of data exposed to 7% noise pollution. This comparative visualization effectively underscores the impact of noise on the data distribution.

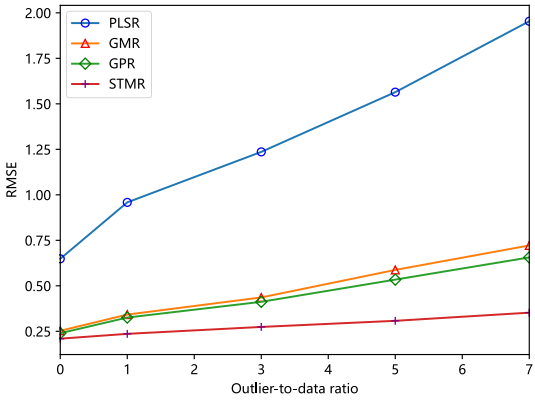

Figure 3 RMSE values of the model at different proportions of outliers

Figure 3 offers a comprehensive evaluation, employing the Root Mean Square Error (RMSE) metrics, to compare the performance of the four models in the presence of varying proportions of outliers. A conspicuous trend unfolds within the visual representation: the STMR model consistently exhibits superior performance as outliers are progressively introduced. As the dataset incrementally incorporates outliers, a discernible yet moderate increase in the RMSE curves of all four models is observed. However, in this context, the performance of the remaining three models experiences a swift degradation. Specifically, both the Gaussian Process Regression (GPR) model and the Gaussian Mixture Regression (GMR) model suffer diminishing accuracy, with the RMSE values of the Gaussian Mixture Regression (GMR) model escalating by 185%, and those of the Gaussian Process Regression (GPR) model increasing by 173% as outlier instances grow to 7% of the dataset. This escalation directly correlates with the

proliferation of outlier instances. Conspicuously, the Partial Least Squares Regression (PLSR) model emerges as particularly vulnerable to the disruptive influence of outliers, evidenced by a substantial 201% elevation in RMSE values. In stark contrast, the STMR model exhibits remarkable resilience, registering a mere 68% increase in RMSE values. This striking contrast underscores the STMR model's exceptional robustness, rendering it minimally susceptible to the adverse influence of outliers. In this regard, the STMR model attains a commendable status, epitomizing both superior predictive accuracy and unparalleled resilience in the face of outlier-influenced datasets.

4.2. Actual industrial data

In this study, the actual data from the #1 ethylene cracker of a large petrochemical company, which has 11 crackers, numbered H-110 to H-120, are used, and each cracker has 8 observation windows, each of which can observe 12 or 13 furnace tubes. The ethylene crackers H-114 and H-115 were selected as the experimental subjects, and the corresponding operating data and the labeled values of the degree of intersection of all the furnace tubes in one operating cycle of the crackers were used as the experimental data to validate the feasibility and validity of the soft measurement algorithm for coking diagnosis with Bayesian t-distributed mixed regression model (STMR). It was also analyzed in comparison with the Gaussian mixture regression model (GMR), Gaussian process regression model (GPR), and partial least squares regression (PLSR). Performance is evaluated using root mean square error (RMSE). A total of six auxiliary variables were screened by the steam cracking process mechanism and the experience of the field experts, which were the outlet temperature of the furnace tube, the inlet temperature of the furnace tube, the outer surface temperature of the furnace tube, the adiabatic pressure ratio, the pressure across the section, and the venturi pressure. The coking degree of the stovepipe was also categorized into four classes. Technicians usually categorize the degree of coking of the furnace tube into four classes: normal, mild coking, moderate coking, and severe coking.

In order to make the output results of the model more accurately express the coking degree of the furnace tube, the output results of the model are processed as follows

$$\hat{y}_{new} = \begin{cases} j_n, & |\hat{y}_{old} - j_n| < 0.5 \\ j_{n+1}, & \hat{y}_{old} - j_n = 0.5 \end{cases} \tag{19}$$

where j_n is the coking degree level of the furnace tube, $\hat{y}_{old}$ is the unprocessed model output, and $\hat{y}_{new}$ is the predicted coking degree level of the furnace tube.

The total number of experimental data samples in this paper is 5000. The number of samples in the training set is 2000 and the number of samples in the prediction set is 3000. the root mean square error RMSE is used as an index to measure the prediction accuracy of the STMR model and other models.

Table 2. Model prediction results

model	RMSE
PLSR	1.1795
GPR	0.3428
GMR	0.2986
STMR	0.1112

The graph of prediction results of the PLSR model, GPR model, GMR model, and STMR model for industrial data is given in Table 2, STMR model has the best model

performance with the smallest RMSE value compared to the other three models. The error distribution of the STMR model is reduced by 90.6%, 67.5%, and 62.8% compared to the PLSR model, GPR model, and GTMR model. The main reason for this analysis is that the STMR model has better robust performance as the model is less affected when dealing with experimental data containing noise.

5. Conclusion

Focusing on the operational data from cracker furnaces, known for their distinctive attributes of multimodality, nonlinearity, non-Gaussianity, and pronounced noise, this study presents a novel approach for coking diagnosis through a Bayesian t-distribution mixed regression (STMR) model. The model's parameter estimation is meticulously conducted within the Bayesian framework, accounting for these intricate characteristics. By means of simulation experiments and real-world industrial data validation, the study establishes that the proposed model not only exhibits remarkable predictive accuracy but also displays robustness in the presence of noise. As a prospect for future investigations, the research aims to explore the amalgamation of deep learning techniques with STMR models. This amalgamation aspires to augment the soft measurement model across dimensions such as characterization, adaptability, and predictive performance. Consequently, this initiative seeks to fortify the model's capability to adeptly address the intricate intricacies inherent in industrial processes linked to cracking and coking.

References

[1] Chen WQ. Research on quantitative evaluation of carburization damage and magnetic detection technology of ethylene cracker furnace tube [D]. Beijing University of Chemical Technology [2023-07-26].
[2] Xiong SH. Simulation and control of coke clearing in tubular ethylene cracker [D]. Beijing: Beijing University of Chemical Technology, 2016.
[3] Wang DW. Causes of coking and prevention and control technology of ethylene cracking furnace[J]. Chemical Management, 2018 (2): 99-99.
[4] Jin PB. Analysis of the impact of high temperature and complex environment on the performance of ethylene cracker tube [D]. China University of Mining and Technology, 2017.
[5] Li M. Effect of chemical cleaning on the operating condition of cracking furnace[J]. Refining and Chemical Industry, 2020, 31(3): 27-29.
[6] Valus M G, Fontoura D V R, Serfaty R, et al. Computational fluid dynamic model for the estimation of coke formation and gas generation inside petrochemical furnace pipes with the use of a kinetic net[J]. The Canadian Journal of Chemical Engineering, 2017, 95(12): 2286-2292.
[7] Symoens S H, Olahova N, Muñoz Gandarillas A E, et al. State-of-the-art of coke formation during steam cracking: Anti-coking surface technologies[J]. Industrial & Engineering Chemistry Research, 2018, 57(48): 16117-16136.
[8] Xin S, et al. Prediction of coke yield of FCC unit using different artificial neural network models[J]. China Pet. Process. Petrochem. Technol, 2016, 18: 102-109.
[9] Karimi H, Olayiwola B, Farag H, et al. Modelling coke formation in an industrial ethane‐cracking furnace for ethylene production[J]. The Canadian Journal of Chemical Engineering, 2020, 98(1): 158-171.
[10] Barza A, Mehri B, Pirouzfar V. Mathematical modeling of ethane cracking furnace of olefin plant with coke formation approach[J]. International Journal of Chemical Reactor Engineering, 2018, 16(9): 20170243.
[11] Abbasali S M Z, Farsi M, Rahimpour M R. Simulation and dynamic optimization of an industrial naphtha thermal cracking furnace based on time variant feeding policy[J]. Chemical Product and Process Modeling, 2017, 13(2): 20170032.

[12] Zhao T, Li P, Cao J. Soft sensor modeling of chemical process based on self-organizing recurrent interval type-2 fuzzy neural network[J]. ISA transactions, 2019, 84: 237-246.
[13] Xin S, Yingya W, Huajian P, et al. Prediction of coke yield of FCC unit using different artificial neural network models[J]. China Pet. Process. Petrochem. Technol, 2016, 18: 102-109.
[14] Chen DK, Wan X, Zhao Z. COT stochastic distribution system modeling of coking process in ethylene cracker based on working condition identification[J]. Chemical Automation and Instrumentation, 2022.
[15] Davis J W. Statistical pattern recognition[J]. 2003.
[16] WEI Meng, LI Jiabo, YE Min, et al. SOC estimation of lithium-ion battery based on Gaussian hybrid regression[J]. Energy Storage Science and Technology, 2020, 9(3): 958.
[17] Yang F, Qiao TZ, Pang YS, et al. Soft measurement of ball mill material level based on improved Gaussian mixture regression[J]. Modern Electronic Technology, 2018, 41(5): 153-158.
[18] Biernacki C, Celeux G, Govaert G. Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models[J]. Computational Statistics & Data Analysis, 2003, 41(3-4): 561-575.
[19] Wei X, Li C. The infinite Student's t-mixture for robust modeling[J]. Signal Processing, 2012, 92(1): 224-234.